DIW Berlin

German Institute
for Economic Research

Discussion Papers

580

Anzelika Zaiceva

Self-Selection and the Returns to Geographic Mobility: What Can Be Learned from the German Reunification "Experiment"

Berlin, May 2006

Opinions expressed in this paper are those of the author and do not necessarily reflect views of the Institute.

IMPRESSUM
© DIW Berlin, 2006
DIW Berlin
German Institute for Economic Research
Königin-Luise-Str. 5
14195 Berlin
Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
www.diw.de

ISSN print edition 1433-0210
ISSN electronic edition 1619-4535

All rights reserved.
Reproduction and distribution
in any form, also in parts,
requires the express written
permission of DIW Berlin.

Self-selection and the returns to geographic mobility: what can be learned from the German reunification "experiment" *

Anzelika Zaiceva[†]

May, 2006

Abstract

This paper investigates the causal effect of geographic labour mobility on income. The returns to German East-West migration and commuting are estimated exploiting the structure of centrally planned economies and a "natural experiment" of German reunification for identification. I find that migration premium is insignificantly different from zero, the returns for commuters equal to four percent of the mean of the total income, and the local average treatment effects for compliers are insignificant. In addition, estimation results suggest no positive self-selection for migrants, and some evidence of positive self-selection for commuters. Based on these results, moving West does not appear to be a highly rewarded option in Germany.

JEL Classification: F22, J61, J62, R23.

Keywords: returns to geographic mobility, causality, treatment effects

---

*I am grateful to Andrea Ichino for his extremely helpful advices and suggestions, to Frank Vella for his econometric guidance and help with the codes, and to Katharina Spiess for granting me access to the geo-code data. I benefited a lot from the discussions at the DIW SOEP group during my stay there as a guest researcher. I also would like to thank Barry Chiswick, Jennifer Hunt, Joachim Frick, Andre Meier, Holger Bonin and seminar participants at the European University Institute, IZA Summer School, Second IZA AM2 Meeting, and IV BRUCCHI LUCHINO Labour economics workshop for their comments. All errors are mine.

[†]European University Institute, Department of Economics. E-mail: anzelika.zaiceva@iue.it

1 Introduction

With cumulative net migration of 7.5% of the original population over the period 1989-2001, East Germany shows the second highest emigration rates (after Albania) among the countries formerly behind the Iron Curtain (Brücker and Trübswetter, 2004, Heiland, 2004). This is, however, much less than was expected, given the similar cultural background between East and West Germany, and some explanations have been suggested in the literature.[1] The emigration rates have tended to slightly increase again since 1997, and there seems to be no sign of the income convergence from 1995 onwards (Figure 1 and OECD, 2001). Moreover, due to the particular geography of Germany, commuting to the West is a popular option for those who do not want to incur fixed costs of moving, and it may substitute for emigration (Hunt, 2000). These phenomena have raised concerns that individuals with high abilities move to the West ("brain drain") and contribute to sluggish economic growth in the East, as well as the question of how large the mobility premium is in the West. These issues are also gaining general importance in light of the eastern enlargement of the European Union in May 2004 and subsequent european East-West migration.[2]

While answering these questions, it is important, however, to separate the pure effect of geographic mobility from the effect of confounding factors. The reason why doing this is difficult is, usually, the unavailability of the relevant data and / or credible exclusion

[1] See Burda (1995) for an option value of waiting theory of migration.

[2] See among others Brücker et al (2003), Zaiceva (2004) for the assessment of potential emigration after EU enlargement, and Gilpin et al (2006) for the first study of the impact of this migration on the UK labour market.

restrictions. As a result, in contrast to the research on another sort of investment in human capital - returns to education, there exists no study up do date that estimates the causal effect of geographic mobility on income. This paper attempts to fill this gap. In its main contribution to the literature, it exploits the structure of centrally planned economy of the former German Democratic Republic (GDR) together with the unique event of German reunification to make causal statements about the returns to geographic mobility from East to West Germany, controlling for the potential self-selection on unobservables. This study also contributes to the literature that explains sluggish East-West German migration after unification (Burda, 1995). Finally, it can be also viewed in the broader context of transition economics and recent eastern enlargement of the European Union.

Migration theory (Roy, 1951, Borjas, 1987, Chiswick, 1999) postulates that migrants will be positively selected if the distribution of earnings is more unequal in the destination region than in the origin.[3] There exists a vast empirical literature on migration, in which the authors have investigated the selectivity issue, using standard Heckman's procedure, or have documented the association between migration and income.[4] The majority of the existing empirical studies on East-West German migration address the question of self-selection indirectly.[5] The first study that explicitly addresses this issue is a recent

[3]Chiswick (1999, 2000) shows that Roy's model is a special case of the human capital model of migration (Sjaastad, 1962), in which migration is viewed as an investment in human capital, and that it occurs if the present discounted value of the lifetime income stream in the destination region, net of migration costs, is higher than the one in the source region.

[4]See, for example, the pioneering works of Nakosteen and Zimmer (1980) and Grant and Vanderkamp (1980), as well as more recents studies of Tunali (2000), Yankow (1999, 2003), Rodgers and Rodgers (2000), Bauer et al (2002) and Yashiv (2004). Most recently, Ham, Li and Reagan (2004) have undertaken an attempt to use propensity score matching to estimate the returns to migration within the US, relying, however, on the strong assumption of unconfoundedness (selection on observables).

[5]Burda (1993), Burda et al (1998) analyze individuals' intentions to move West. Hunt (2000) estimates the reduced form multinomial logit of the decisions to move, to commute or to stay. See also Hunt (2006)

paper by Brücker and Trübswetter (2004), in which the authors find no robust evidence of positive self-selection on unobservables for migrants over 1994-1997. As for the mobility premium, Hunt (2001) shows that those who took a job in the West between 1990-1991 enjoyed large wage gains, but that the correlation between income and migration is small or insignificant for the subsequent movers. She concludes that an economy undergoing a successful transition would initially have high returns to moving, which would fall as the transition progressed. It is not clear, however, what kind of effect is estimated[6] and how the selection on unobservables is dealt with.[7] Thus, although the empirical studies have analysed an association between income and migration, they have failed to plausibly identify causal relationship.

This paper exploits programme evaluation techniques, and, using the language of that literature, attempts to identify the effect of treatment (geographic mobility) on the treated (mover), as well as the effect for compliers (a subpopulation of movers whose status changes with the instrument). I investigate this question using both parametric and nonparametric econometric methodologies. I use the GSOEP dataset, which has a longitudinal structure due to which it is possible clearly to identify movers. Another big advantage of this dataset is that it contains pre-unification information. The main disadvantage, however, is the small number of observations for movers. I merge it with the confidential geo-code data on the individual place of residence.

---

and the references therein.

[6]See Heckman, LaLonde and Smith (1999) for a discussion of when it may be more useful to estimate the effect of treatment on the treated than to estimate an average treatment effect.

[7]Hunt (2001) recognizes these issues by noting that "the coefficients on the moving dummies should not be interpreted as the return to an exogenous move by a random worker, because movers are likely to be unobservably different from stayers" (p.192).

The following instruments are constructed for migration and commuting respectively in order to identify the returns to geographic mobility - home ownership before unification, and a dummy which equals one if an individual before unification lived in a county (*"kreis"*) that had a common border with West Germany. Both instruments approximate theoretical costs of moving: the former captures the well-established negative relation between home ownership and the propensity to migrate, while the later captures the costs of commuting West that increase with distance from the border. As shown below, in the former GDR housing decisions and voluntary geographic labour mobility were usually restricted by political considerations. In making these decisions, little was left to the unobservable abilities and motivation. Finally, German reunification was not anticipated by anybody until shortly before the event. Although one still may argue that the allocation of housing, jobs and residence of individuals in the Communist economy was not random, it was mostly based on the factors that are not relevant for the market economy and the post-unification individual incomes, which are, thus, ignorable. Of course, the validity of these instruments is open to doubt; however, the evidence presented in this paper seems to suggest that the assumptions hold at least for commuters.

The main findings of the paper are as follows. First, I find no evidence of positive selection on unobservables for East-West German migrants and some evidence of positive self-selection for commuters. Second, the mobility premium is small for both migrants and commuters, independently of the model used and the assumptions made. Both parametric and nonparametric selection models suggest no significant returns for migrants in terms of total long-term income, and the local average treatment effect for compliers is also insignificantly different from zero. The returns for commuters are slightly higher and equal to 4% of the mean of the total income, and the local average treatment effect for compliers is insignificant.

The paper is organized as follows. Section 2 provides the description of the data, definitions and sample selection. Section 3 justifies the instruments. Section 4 outlines estimation strategy. Estimation results are discussed in section 5, and section 6 provides a sensitivity analysis. Section 7 concludes.

2 Data, definitions and sample selection

The data used in this paper are extracted from the public use file of the representative German panel household survey (GSOEP)[8], and are merged with the confidential geographical coding of individual places of residence. Due to the GSOEP's longitudinal structure, it is possible to identify and trace movers (and their incomes). Another advantage of this dataset is that the first wave of the eastern sample was drawn in June 1990, i.e. before the monetary union and formal unification took place, and, thus, it provides a unique opportunity to use pre-unification data to construct the exogenous source of variation in mobility. The main disadvantage of the dataset, however, is the small number of observations for movers.

An individual is defined as a migrant if he has changed his residence from East to West Germany at least once during 1990-2001; otherwise he is a stayer. An individual is

[8] See SOEP Group (2001) for a description of the dataset.

a commuter if he lives in the East and his region of work is West Germany in any of the years 1990-2001.[9] A definition of income is not trivial in such a study. Theory suggests that while making a decision to move, an individual takes into account his total lifetime income, and empirical studies find that the assimilation period matters.[10] In order to be consistent with the theoretical definition of lifetime income, as well as willing to avoid the problem of transitory income drop right after move and to save observations, I have used the mean of annual incomes as a dependent variable.[11] I thus average over the available years for stayers. For migrants, I average over the available years after an individual migrates, and for commuters - over the years during which an individual commutes. The total annual income is defined as a sum of labour income (sum of wages, second-job and self-employment earnings) and various social security benefits (such as unemployment benefits, maternity benefits etc). The mean income is set to missing only if information on all the components is missing.[12] All incomes are inflated to 2001 by regional CPIs and are expressed in DM.

The instruments used in this study are as follows. For migration, I construct a dummy

[9]Note that by defining migrants in this way I have to include commuters within "stayers", and by defining commuters, I have to include actual and potential migrants within "stayers". Although dropping them from a control group could help to clarify the definitions, this would introduce a sample selection problem. Nevertheless, below I also experiment with excluding actual and potential movers from the respective comparison groups (see Section 6).

[10]It is argued that estimates based on earnings data with limited time horizons will not capture life-cycle wage growth, tending to downward bias in the estimated returns (Greenwood, 1997). See Yankow (2003), Greenwood (1997) for a discussion of the potential biases, and Yankow (1999) and Rogers and Rogers (2000) for the attempts to capture the long-term earnings effects.

[11]Siebern (2000), Carneiro and Heckman (2002), and Carneiro and Lee (2004) use similar cumulative definitions.

[12]I also exclude the obvious outliers from the sample, i.e. individuals whose average annual income is less than 1000 DM (19 observations) or greater than 130000 (5 observations). I have experimented with the lower threshold at 100, 500 and 1000 DM, and the upper threshold at 100 000 DM, using the so-called "winsorising" procedure, in which 2.5% of the outliers from both tails were given the closest neighbour's value, and keeping all individuals in the sample. The results were not affected to any great extent.

which equals one if an individual was a home-owner in 1990, and is zero otherwise. 32% of respondents in 1990 in East Germany reported being a house / flat owners. For commuting, an instrument equals one if an individual resided in a county ("kreise") close to the Western border (i.e. that had a common border with West Germany or West Berlin) before unification.[13] Approximately 30% of persons lived in such counties in 1990.

I restrict the sample to easterners who were living in East Germany in 1990, exclude pensioners and students, and use the incomes of individuals who are 18 years old or more in each year.[14] Final sample sizes in the most restricted specifications are 3,043 observations for migration (of whom around 6% are migrants), and 2,953 observations for commuting (of whom around 15% are commuters). In line with the aggregate data on migration in Figure 1, the number of migrants was large right after unification, then decreased, but tends to increase again since 1997. Commuting seems to follow an opposite trend.

Kernel densities of average total annual incomes for movers and stayers are shown in Figure 2. As expected, the distribution of incomes for stayers is more compressed, and there are more migrants and commuters in the upper tail of income distribution. Descriptive statistics for the key variables is given in Table A1 in the Appendix. The first two columns show means and standard deviations for migrants and stayers, the last two

[13] One may suggest dropping Berlin from the sample given its very specific geographical and political situation. I reestimated all the models below dropping Berlin, and the results did not change much, the main difference being a smaller effect for commuters. These results are available upon request.

[14] I also drop the return and multiple migrants (around 20%) and do not analyze them separately in this paper, given the insufficient number of observations, unclear lifetime income definition, and since I am interested in the returns to permanent migration. See, for example, Rogers and Rogers (2000) for a similar approach. Recognizing, however, that the exclusion of multiple movers could bias the results, in the sensitivity analysis below I also estimate the models using data for all migrants.

- for commuters and stayers. As can be seen from the table, all potential movers have on average a higher total annual income than stayers. Compared to stayers, migrants tend not to own a house in 1990, and commuters tend to live in the border regions in 1990. As expected, potential movers are younger, single and better educated than stayers. Thus, there exists a preliminary evidence of positive selection of both migrants and commuters with regard to pre-treatment university degree and income. There are more males among commuters, however, more females among migrants. On the other hand, there are fewer individuals with any kind of vocational training among any movers, more government sector employees among migrants, but fewer among commuters, and fewer blue-collar employees among migrants, but more among commuters. Table A1 presents some systematic differences in observable characteristics between movers and stayers; thus, there is reason to suspect, a priori, that selection on unobservables will be an issue. To cope with this, I rely in the remainder of the paper on the instrumental variables, which are justified in the next section.

3 Are the instruments legitimate?

In order to make causal statements about the returns to geographic mobility, it is important to justify the validity of the instruments. Unfortunately, this assumption cannot be tested, and one has to rely on the available general facts. To be a valid instrument, pre-unification home ownership and residence dummies must affect income only through migration or commuting, i.e. they must be uncorrelated with any non-ignorable confounding factors that affect ex-post income in the market economy, such as ability or motivation. This can be justified by referring to the structure of centrally planned economies.[15]

In GDR, as in any communist societies, there was a high degree of centralization in the labour and product markets: all firms were owned by the state and an elaborate plan directed the allocation of inputs, the distribution of outputs, wage levels and prices (Krueger and Pischke, 1995)[16]. To secure constant prices for inhabitants, the state bore 80% of costs of basic supplies, from bread to housing. Shortages were a norm. The distribution of income was compressed, and wage inequality, as measured by the Gini coefficient, was very low.[17] Official unemployment was absent, since workers were kept inefficiently in companies even if they were unproductive, or the government quickly found a new job for anybody who might have been displaced in order to achieve the goal of full employment.[18] The communist ideology stressed uniformity of outcomes, irrespective of individual differences in ability or effort. Political tolerance was also important: the system functioned smoothly only when its component parts were staffed with individuals whose values coincided with those of the regime.

Moreover, housing and occupational choices, and thus voluntary geographic labour mobility, were restricted. In principle, everyone had a right to a house; however, due

[15] Fuchs-Schündeln and Schündeln (2005) use a similar logic and the structure of the centrally planned economy of ex-GDR to identify precautionary savings in post-unification East Germany.

[16] In 1985, for example, state-owned enterprises and collectives earned 96.7 percent of the total net national income.

[17] Fuchs-Schündeln and Schündeln (2005) report that in 1988, the average net income of individuals with a university degree was only 15% higher than that of blue-collar workers. Also, intersectoral differences in net incomes were minimal, amounting to only 150 Marks per month on average with an average monthly income of around 1,100 Marks in 1988.

[18] The significant misallocation of labour in centrally planned economies is well known. See, for instance, Kruger and Pischke (1995) for a comparison of East and West German labour markets before and after unification.

to rationing by the state (i.e. a mechanism called the System of Material Balances), long queues were a norm.[19] Access to housing quality was regulated largely through informal (and often politically mediated) networks (Buechtemann and Schupp, 1992). In many ways access to material and social activities in ex-GDR was mediated through the sphere of work, and, in particular, the FDGB unions acted as the prime political links between the working population and the Socialist power elite embodied in the SED party, and as key agents in the distribution of housing (ibid). In general, flats may have been allocated to individuals due to urgent need or merit, personal connections or corruption, or by inheritance. And those who paid a nominal rent for a state-owned flat enjoyed considerable consumer surplus (Kornai, 1980). As for occupational choice, job offers were usually made to individuals right after the completion of their education and according to the Socialist plan. Even admissions to the various fields were regulated by the plan.[20]

Overall, the Communist system operated like a large internal labour market, with rules and party membership playing an important role in the allocation of jobs and wages (Krueger and Pischke, 1995). Thus, little was left to the individual abilities and motivation. Finally, the fall of the Berlin Wall in 1989 could not been foreseen. Therefore, to the extent that individuals had not been self-selecting into home ownership statuses or into the regions on the basis of their unobservable characteristics relevant for the market economy, the instruments provide the exogenous source of variation in mobility, and the

[19]The "waiting list" could be very long. For example, the wait for an apartment in the Soviet Union during the 1980s was typically 10 to 15 years; as a result, families had to plan and buy the housing for their children to live in in advance (www.wikipedia.com).

[20]In fact, only a certain quota of students was allowed to complete the last two years of high school, necessary to attend university. Additional criteria were membership in the official youth organisation, political tolerance, and family background (Fuchs-Schündeln and Schündeln, 2005).

assignment to treatment is strongly ignorable.

However, the exclusion restriction assumption is violated if pre-unification home ownership directly affects post-unification income of migrants (after controlling for human capital, occupation and regional macro-indicators), and if living in the border regions or not before unification per se matters for the ex-post income of commuters (after controlling for the same factors). The former would be true if, for example, more able persons were also more successful in gaining access to their own housing, leading to an upward bias in the estimates. As for the latter, in the former GDR, it was likely for only those who supported the regime (i.e. party members and the so-called "nomenklatura") to be allowed to live close to the western border. If these people were also more motivated / active / able, the validity of the instrument may be violated unless one controls for the "nomenklatura effect". Fortunately, Bird, Frick and Wagner (1998) provide a proxy for party membership and nomenklatura status - telephone availability before unification.[21] Thus, in Section 6 I also control for this effect.

Finally, before using different models to estimate the effect of moving, an informal exercise can be undertaken in order to further justify the instruments. If the instruments approximate a randomized experiment, the characteristics of those for whom the instrument equals one must be equal to those for whom it equals zero, meaning that persons are randomly assigned across the two groups. Table A2 shows that for migration, the home ownership dummy is indeed orthogonal to some covariates, although some differences (at

[21] They argue that in GDR only 20% of individuals had a telephone before unification (in West Germany, 97%), and a telephone was a sort of in-kind benefit for distinguished party members, since, like many other goods, it was not freely available in the economy of shortages.

5%) seem to exist in many of them. In contrast to the expectations, however, the more educated and those having a higher pre-treatment individual income are *less* likely to own a house before unification. Thus, it is likely that housing was not randomly allocated to individuals in the Communist economy, however such allocation was probably based on some political factors and personal connections and not on the unobservables that are relevant for the market economy, such as individual ability. Moreover, differences in most characteristics, although statistically significant, are not large economically, i.e. differences in all characteristics range from 9 to 20% of the respective standard deviations. For commuting, the border regions dummy is orthogonal to all covariates with the exception of telephone availability in 1990, which actually confirms the fact that only politically tolerant persons lived / were allowed to live in the border regions.

Therefore, although one still may argue that the allocation of housing, jobs and residence of individuals in the Communist economy was not random, it was mostly based on the factors that are not relevant for the market economy and the post-unification individual incomes. Thus, I believe that the evidence presented in this section allows us to use the two instruments for the analysis and to make a valid causal inference, at least for commuting.

4 Econometric methodology

In order to estimate the causal effect of geographic mobility on the income of movers, potential outcomes model is used. Let Y_{1i} and Y_{0i} denote individual i's potential income with and without movement. Then:

$$Y_{1i} = X_{1i}\beta_1 + \varepsilon_{1i} \quad (1)$$

$$Y_{0i} = X_{0i}\beta_0 + \varepsilon_{0i} \quad (2)$$

where X_{ki} are individual socio-economic characteristics, $\beta's$ are unknown parameters, $E(\varepsilon_{ki}) = 0$, and $k = \{0, 1\}$. Let $D_i = 1$ if individual is a mover, and $D_i = 0$ otherwise. The outcome is observed only in one state, i.e. $Y_i(D_i) = D_iY_{1i} + (1 - D_i)Y_{0i}$. After some manipulations one can derive the following model:

$$Y_i = \alpha_0 + X_i\beta + \Delta_i D_i + \eta_i \quad (3)$$

where Y_i is the observed outcome, and the "unconditional" error term η_i has a zero mean, i.e. $E(\eta_i|X_i) = 0$, but $E(\eta_i|D_i, X_i) \neq 0$.[22]

Assuming further that there exist costs of moving C_i, the following selection rule applies:

$$D_i = I(Y_{1i} - Y_{0i} - C_i > 0) = I(Z_i\gamma + u_i > 0) \quad (4)$$

where Z_i is a vector of exogenous variables, γ are the reduced form parameters and $E(u_i) = 0$. The errors $(\varepsilon_{1i}, \varepsilon_{0i}, u_i)$ are assumed to be correlated with covariances σ_{ki},

[22] This is a so-called "random coefficients model", in which there are potentially two sources of unobserved heterogeneity: one that influences both the decision to move and labour market outcomes of individuals (heterogeneity in η_i), and another that is related to the idiosyncratic gain from mobility (heterogeneity in responses to treatment Δ_i).

$k = \{0, 1\}$. The self-selection works through this correlation in the errors.

The effect of interest in this study is an average effect of treatment on the treated (ATT). Formally it can be written as follows:

$$
\begin{aligned}
ATT &= E(\Delta_i | Z_i, D_i = 1) = E(Y_{1i} - Y_{0i} | Z_i, D_i = 1) = \\
&= E(Y_{1i} | Z_i, D_i = 1) - E(Y_{0i} | Z_i, D_i = 1) = \\
&= E(\Delta_i) + E(\eta_i | Z_i, D_i = 1) \qquad (5)
\end{aligned}
$$

where the effect is the difference between actual outcome for movers and a counterfactual outcome for movers had they stayed. It equals to the average effect for a random person in the population *plus* the idiosyncratic gain from treatment (the returns to unobservables), and there is no a priori reason to expect $E(\eta_i | Z_i, D_i = 1) = 0$. Thus, the OLS estimation of (5) provides biased and inconsistent estimates.

To estimate the effect of moving West on income I, first, estimate parametric sample selection model of Heckman (1976, 1979), then - the nonparametric sample selection model of Das, Newey and Vella (2003), and calculate the treatment effect as the difference between the actual outcome for movers and the counterfactual outcome for movers have they stayed. I then also estimate the local average treatment effect for compliers (Angrist, Imbens and Rubin, 1996).

The specification most widely used in migration studies is related to the representative agent model and is estimated by the two-step parametric procedure developed by Heckman (1976, 1979), wich assumes no idiosyncratic gain from treatment. The correction function is the inverse Mill's ratio for each subsample. Note that this procedure requires exclusion restrictons. In addition, if the joint normality assumption does not hold, it will produce inconsistent estimates.

The nonparametric sample selection model that does not impose any distributional assumptions and does not restrict the form of the correction function allows to overcome the disadvantages of the above mentioned parametric approach. The estimation of such model is considered in Das et al (2003), building on the prior work of Newey (1988). The identification requires exclusion restrictions, and the model is identified up to an additive constant. The approach amounts to estimating in the first step a conditional probability of selection (propensity score) without making any distributional assumptions, and, in the second step, to approximating the correction function with polynomial series. The order of the correction term can be chosen using a leave-one-out cross-validation criterion.[23] For the purposes of this paper, the estimation of the intercept is crucial. Hence, I also use two semiparametric techniques to consistently estimate the intercept (Heckman, 1990 and Andrews and Schafgans, 1998).[24]

Finally, making no restrictions on unobserved heterogeneity and no distributional assumptions, the Local Average Treatment Effect (LATE) - causal treatment effect for the subpopulation of compliers, can be estimated (Angrist et al, 1996).[25] LATE has been

[23] This is the sum of squares of forecast errors, where all the other observations were used to predict each single observation, and the specification with the smallest sum of forecast errors is chosen.

[24] I use 50% of both subsamples as a threshold value.

[25] Note that the Random Assignment, Exclusion Restrictions and a Non-zero Effect of the Instrument on the Treatment assumptions are satisfied based on the evidence presented in Sections 3 and 5, respectively. Stable Unit Treatment Value assumption seems plausible, since movers constitute only a small fraction of the population, thus ruling out general equilibrium effects. Finally, the assumption of Monotonicity (no defiers) also seems plausible, since both owing a house and living far from the border constitute costs for mobility.

criticised for two reasons: it is identified only for a small fraction of population, which is unobservable, and it is also instrument-dependent.

5 Estimating the effect of mobility on income

I use the standard Mincerian specification of the income functions. Variables such as experience, education and marital status in 2001 are endogenous; thus, in my preferred specification, I use only exogenous variables, such as sex, age and its square (as a proxy for experience) and the predetermined marital status (as a proxy for "psychic" migration costs) and human capital variables in 1990.

As can be seen from the second last row of Table A3, the instruments are "strong".[26] As expected and consistent with other studies, the correlation between pre-unification home ownership and propensity to migrate is negative, reflecting theoretical costs of migration. The border with the West dummy has a large positive impact on the probability of commuting and indicates that the costs of commuting, indeed increase with distance.

5.1 Results for migration

Assuming no idiosyncratic gain from migration and willing to compare my results to the existent literature, I first estimate the standard Heckman's selection model. The first stage probit estimates (Table A3 column 1) confirm that, on average, younger people are more likely to move West, and home owners are less likely to migrate, consistent with

[26] As a rule of thumb, to be considered "strong", the F-statistics from the linear regression of the endogenous variable on the instrument must be not less than 10 (see Stock, Wright and Yogo, 2002).

expectations and in line with previous studies.[27] Probit marginal effects indicate that an additional year decreases probability of moving by 0.2 percentage points, while owing a house decreases the likelihood of moving West by 4 percentage points. Males are 1 percentage point less likely to move West. Both university degree and marital status have expected signs, but neither these variables nor vocational education or government sector employment are significant. Finally, neither blue collar occupation in 1990, nor the state's unemployment rate affects probability to move West. Heckman's second stage estimates for migrants[28] (Table A4 column 1) suggest that males have a higher total income than females, experience as proxied by age and its square has the traditional concave profile, and university graduates earn more. However, neither vocational education nor occupational dummies are significant for movers, suggesting that partly human capital acquired in the centrally planned economy is not transferable / valuable in the West. Being married in 1990 reduces the ex-post income of movers. The coefficient on the inverse Mills ratio is positive, but insignificant, thus indicating no evidence of positive self-selection for movers, which is partly consistent with Brücker and Trübswetter (2004). Estimates for stayers (Table A4 column 2) suggest that, on average, male stayers have a higher total income than females, university graduates earn more, experience has the expected sign, those who had a vocational degree and were employed in the government sector in 1990 earn more, and those in blue-collar occupations in 1990 earn less. The Mills ratio for stayers is also insignificant, which is in contrast to Brücker and Trübswetter (2004), where the negative

[27] Note, however, that when the age squared is added to the probit regression, both age variables become insignificant.

[28] Standard errors in the second stage are corrected both for heteroskedasticity and generated regressors (see Greene, 1981, Newey, 1984, Pagan, 1984).

sign on the Mills ratio was found across 1994-1997.[29] Finally, to test the normality assumption I use the conditional moment test[30], which indicates that normality cannot be rejected, implying that Heckman's estimates are consistent.

Nevertheless, I also experiment with the nonparametric sample selection model. In the first stage, I estimate linear probability model (see Table A3 column 2) and construct predicted probabilities. I also trim on propensity scores as suggested in Das et al (2003). The cross-validation criterion indicates the linear correction function for movers and a polynomial of order 3 for stayers (Table A6). Table A5 (columns 1-2) shows the non-parametric second stage estimates.[31] The coefficients on covariates for both stayers and movers are quite similar to the parametric Heckman's model, apart from the correction terms. When normality is not imposed, there is again no evidence of positive self-selection for movers.

Finally, imposing neither distributional assumptions nor restrictions on the unobserved heterogeneity, I estimate the model by IV-LATE framework of Angrist et al (1996). Table A7 summarizes the so-called intention-to-treat effects (reduced form migration and income equations), structural IV and OLS estimates (upper panel A). Column 1 shows the coefficient of the home ownership dummy in regressions for migration, and column 2 - the coefficients of this dummy in the reduced form income equations (i.e. models

[29]However, one needs to keep in mind that the time span in this study is much larger, the definitions of migrants and income are different, and the exclusion restrictions are also different.

[30]See Newey (1985), Pagan and Vella (1989). To carry out the test, I construct the relevant moment conditions (3rd and 4th moments) and regress them on a constant and scores from probit. Standard errors on constants indicate whether normality holds.

[31]Standard errors are calculated according to the variance-covariance formula in Das et al (2003) and are corrected for both heteroskedasticity and generated regressors.

that exclude migration). Column 3 reports the IV estimates of the return to migration, which are the ratios of the corresponding intentions-to-treat effects, and OLS estimates are shown in column 4 for comparative purposes. As can be seen from this table, the IV coefficient is not statistically significant. The standard errors are traditionally very large, and the difference between the OLS and IV could be due to the sampling error. Thus, the local average treatment effect for compliers shows that those individuals who migrate if they did not own a house in 1990, and would have not migrated if they had owned a house, have no significant returns to their ex-post total long-run income from migration. One should bear in mind, however, that the results for migration have to be interpreted with caution: there might still exist some doubts on the validity of the instrument, and standard errors of the IV point estimate are very large flipping from large negative to positive.

Table A8 shows treatment effects for migrants in the different econometric models used.[32] Both parametric and nonparametric selection models produce not statistically significant effects.[33] Therefore, the effect of migration for migrants is not statistically different from zero, and the local average treatment effect for a subpopulation of compliers is also not significant.[34]

Overall, several interesting findings occur from the estimates. First, no evidence of positive self-selection on unobservables for East-West German migrants during 1990-2001

[32] To test the null of no significance of treatment effects for sample selection models, the t-statistics is calculated as in the Oaxaca decomposition (Greene, 2000).

[33] at 5% level.

[34] Restricted model without human capital covariates generated qualitatively identical results (available upon request).

is found. Such a result is partly in line with Brücker and Trübswetter (2004), and is also consistent with the theoretical predictions of the human capital model on self-selection (Chiswick, 1999), when direct out-of-pocket costs of migration are small. Given that the inequality of earnings in East Germany has approached West German levels in the late 1990s, the standard Roy's model would also predict that a positive selection bias of East-West migrants should disappear. This, however, can also result from the above-mentioned definition of migration: the most able may choose to migrate, but also to commute or to stay because of the new opportunities in the East, thus resulting in no significant self-selection. Or it may also be due to the aggregation of data over ten years, i.e. the cohort quality effect might be at work here, the first migrants being of better quality than the subsequent movers.[35] Finally, also a small number of observations for movers may play a role.[36]

Second, both the treatment effect for migrants and the local average treatment effect for compliers are insignificant.[37] This result might be a consequence of high unemployment in the East, when people move West not in search of a higher income, but to escape from unemployment, and it may also be the cause of return migration to the East. Together with no positive selection for migrants, it may also reflect attitudes towards risk[38], or

[35]Ideally, one would need to run year-by-year regressions; unfortunately, the very small number of observations for migrants preclude me from doing this. I have, however, reestimated all the models for two periods, 1990-1995 and 1996-2001, and the results are available upon request.

[36]Note, however that the inverse Mills ratio is also insignificant in the restricted model without lagged education and occupation dummies.

[37]Given significant and positive OLS effect and insignificant effect in Heckman's and IV models, insignificant Mills ratio are somewhat surprising. As mentioned above, this can be a consequence of the small number of observations for movers. The upward bias in OLS can also be due to the omitted variables, where such variables are other (observable) characteristics that are positively corellated with the regressors.

[38]Dohmen et al (2005) find that the East-West migrants are more willing to take risks in general (thus

non-transferable human capital. Finally, the exclusion of earlier migration (1989-1990) from the analysis due to the unavailability of data may bias the effects downward, since high initial migration most probably left behind those with the highest migration costs.

These results, however, are not entirely surprising. Hunt (2001) finds that the wage gain is small or insignificant for the post-1991 movers. Burda and Hunt (2001) find that increased by western unions wages, indeed, kept the East Germans at home. Parikh and Van Leuvensteijn (2003) conclude that because of wage convergence between East and West Germany, wages did not play any consistent role in determining migration over 1991-1999. A relevant question to ask then is, why people have decided to move if there are no returns to migration. Possible answers could include too optimistic expectations while calculating present value of the future income streams, other non-economic (for example, ideological) reasons, or chosing an option to commute instead. Overall small number of East-West German migrants and return migration to the East is, in fact, in line with these findings.

5.2 Results for commuters

In order to estimate the treatment effects for commuters, I follow the same procedures as with migrants. Reduced form probit estimates (Table A3 column 3) suggest that on average males, young and those with university degrees and living in the border regions in 1990 are more likely to commute West. Interestingly, blue-collar workers also have

are relatively less risk averse), however, they are less willing to take risks in their careers (and are thus more risk averse with respect to their labour market incomes).

a higher probability of commuting. And, as expected, individuals from the disadvantaged states tend to commute more. Second-stage parametric Heckman's estimates for commuters (Table A4 column 3) suggest that males and university graduates earn more (although much less than migrants, suggesting again that human capital acquired in the communist economy is only partially transferred to the West), and that experience has a traditional concave profile. For stayers (Table A4 column 4), in addition, being employed in the government sector and having a vocational degree in 1990 affect their ex-post incomes positively, while being a blue-collar employee in 1990 affects this negatively. The selection correction terms are insignificant for both commuters and stayers. However, the conditional moment test rejects the normality assumption, implying that parametric estimates are inconsistent.

Estimates of the linear probability model, which are used to construct propensity scores for the nonparametric sample selection model, are shown in Table A3 column 4. The leave-one-out cross-validation criterion (Table A6) suggests a polynomial of order 2 for commuters and no correction polinomial for stayers. The estimated coefficients for both commuters and stayers (Table A5 columns 3-4) are again similar to those in the parametric model, apart from the correction terms. In addition, the marginal effects of the correction functions for commuters are positive, suggesting, thus, positive self-selection for commuters.

Finally, I estimate the IV-LATE model of Angrist et al (1996), relaxing all the distributional assumptions and assumptions of homogeneity. The lower panel B of Table A7 shows the respective intentions-to-treat effects, structural IV and OLS estimates. Again, IV point estimates are not statistically significant. Thus, the local average treatment effect for persons who commute if they were living in the border regions in 1990 and who would not have commuted otherwise, is not statistically different from zero.

Table A8 shows the effects for commuters in different econometric models used.[39] Overall, for commuters, positive self-selection seems to be present. The local average treatment effect for a subpopulation of compliers is again insignificant. The treatment effects for all commuters, however, are equal to 4% of the mean of the total income (which approximately equals ten). This is, in fact, what was expected, since migration usually involves higher costs than commuting, indicating that the overall returns should be higher for the latter.

6 Robustness checks

In addition to the changes in specification reported above, the following sensitivity analysis was undertaken. First, I check how robust the results are to the inclusion of additional controls. I include a dummy which equals one if a person was unemployed in 1990 to check how the lagged employment status influences both decision to move and ex-post incomes. I then add the household monthly income in 1990 in order to capture additional household-level characteristics. Second, I improve on the validity of the instruments controlling for the "nomenklatura" effect mentioned above. I use a telephone availability in 1990 dummy to control for this social background. One may also argue that apart the

[39] Again, restricted model without human capital covariates generated qualitatively identical results (available upon request).

"nomenklatura" effect, it is also important to control for the ideology, since it may be an important determinant of migration. Thus, to proxy for the ideological reasons I then also include a variable that ranks political interests of a person before unification (from "very strong" to "none"). Third, I exclude the self-employed from the sample, since there might be self-selection into this group. Fourth, I retain all return and multiple movers in the sample. Finally, I improve the definition of the control group: I drop commuters from the control group for migrants, and migrants - from the control group for commuters. It is expected that, since the western wage earners are excluded from the stayers, the effects should be larger.[40] [41]

Table A9 shows these sensitivity checks. In general, the effects are similar to those reported in Table A8. Contrary to expectations, dropping the actual and potential movers from the respective control groups does not change the treatment effects. This, however, contributes to the robustness of the results.[42]

[40] One could also argue that home ownership statuses before unification were different in urban and rural areas, the majority in rural areas being home owners. To control for this, data on rural / urban residence before unification is needed. Such data, however, has not been available to the author.

[41] I have also estimated all the models dropping observations with missing 1990 information from the sample. The results are not affected for commuters, but are not robust for migrants. The overall results are available upon request.

[42] In addition, all models have been reestimated using labour income as a dependent variable. The results for migration were qualitatively the same, OLS delivering the only significant and slightly higher estimates (approx. 4% of the mean of the labour income). For commuters, consistent nonparametric estimates were again significant, but slightly higher (4.6% of the mean of the labour income). Moreover, LATE for compliers was marginally significant and equaled approximately to 4% of the mean of the labour income. The contribution of the selection bias was again the same. These results seem to suggest that commuting particularly pays-off with respect to the labour income, which is, in fact, true by definition of commuters.

7 Concluding remarks

The question of the returns to geographic mobility, especially in the context of transition economies, remains difficult to deal with, mainly due to data availability and identification problems. This paper exploited a structure of the centrally planned economy of ex-GDR and a "natural experiment" of German reunification, and attempted to make a causal inference for the returns to East-West German migration and commuting. Pre-unification home ownership was argued to provide an exogenous source of variation in migration, and proximity to the West German border before unification - in commuting. Both parametric and nonparametric sample selection models were estimated to control for selection bias, and the effects of treatment (geographic mobility) on the treated (movers) were calculated. Further, the local average treatment effect for the subpopulation of compliers was estimated.

The main findings from this study are as follows. First, no evidence of positive selection on unobservables for migrants is found, and positive self-selection for commuters seems to exist. Second, no significant returns to migration in terms of total long-term income were found. The returns for commuters are equal approximately 4% of the mean of the total income; however, they are also insignificant for compliers. A higher overall gain for commuters is in line with expectations, taking into account the higher costs of migration and lower unemployment rate for commuters than for migrants. This result may also suggest that commuting might indeed be a substitute for migration. Third, these findings seem to be robust to different changes in specification and in the sample. Based on these results, in the long run moving West does not appear to be a highly rewarded option for eastern Germans. This fact, although subject to the assumptions and definitions used in this study, could constitute an important part of the explanation of the sluggish East-West migration in Germany.

In addition, one should bear in mind that the results for migration have to be interpreted with caution. Moreover, the multinomial sample selection model, where the choices are to migrate, commute or stay might deliver more precise estimates. This is left, however, for future research.

References

[1] Andrews, Donald W., and Marcia A. Schafgans. 1998. Semiparametric estimation of the intercept of a sample selection model. Review of Economic Studies 65, no. 3:497-517.

[2] Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. Identification of causal effects using instrumental variables. Journal of the American Statistical Association 91:444-72.

[3] Bauer, Thomas, Pedro T. Pereira, Michael Vogler, and Klaus F. Zimmermann. 2002. Portuguese migrants in the German labor market: Performance and self-selection". International Migration Review 36, no. 2:467-91.

[4] Bird, Edward J., Joachim R. Frick, and Gert G. Wagner. 1998. The income of socialist upper class during the transition to capitalism: Evidence from longitudinal East German data. Journal of Comparative Economics 26, no. 2:211-25.

[5] Borjas, George J. 1987. Self-selection and the earnings of immigrants. American Economic Review 77, no. 4:531-53.

[6] Brücker, Herbert, and Parvati Trübswetter. 2004. Do the best go West? An analysis of the self-selection of employed East-West migrants in Germany. Discussion Paper no. 986, Instituts zur Zukunft der Arbeit (IZA), Bonn, Germany.

[7] Buechtemann, Christoph F., and Jürgen Schupp. 1992. Repercussions of reunification: Patterns and trends in the socio-economic transformation of East Germany. Industrial Relations Journal 23, no. 2:90-106.

[8] Burda, Michael C. 1993. The determinants of East-West migration: Some first results. European Economic Review 37: 452-61.

[9] Burda, Michael C. 1995. Migration and the option value of waiting. Economic and Social Review 27: 1-19.

[10] Burda, Michael C., and Jennifer Hunt. 2001. From reunification to economic integration: Productivity and labor markets in Germany. Brookings Papers on Economic Activity 2:1-71.

[11] Burda, Michael C., Wolfgang Härdle, Marlene Müller, and Axel Werwatz. 1998. Semiparametric analysis of German East-West migration intentions: Facts and theory. Journal of Applied Econometrics 13, no. 5:525-41.

[12] Carneiro, Pedro, and James J. Heckman. 2002. The evidence of credit constraints in post-secondary schooling. The Economic Journal 112, no. 482:705-34.

[13] Carneiro, Pedro, and Sokbae Lee. 2004. Comparative advantage and schooling. Working Paper, University College London, London, UK.

[14] Chiswick, Barry R. 1999. Are immigrants favorably self-selected? American Economic Review 89, no. 2:181-5.

[15] Chiswick, Barry R. 2000. Are immigrants favorably self-selected? An economic analysis. In Migration theory: Talking across disciplines, ed. Caroline B. Brettell and James F. Hollified. New York: Routledge.

[16] Das, Mitali, Whitney K. Newey, and Francis G. Vella. 2003. Nonparametric estimation of sample selection models. Review of Economic Studies 70:33-58.

[17] Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner. 2005. Individual risk attitudes: New evidence from a large, representative, experimentally-validated survey. Discussion Paper no. 1730, Instituts zur Zukunft der Arbeit (IZA), Bonn, Germany.

[18] Fuchs-Schündeln, Nicola, and Matthias Schündeln. 2005. Precautionary savings and self-selection - Evidence from the German reunification "experiment". Quarterly Journal of Economics 120, no. 3:1085-1120.

[19] Greene, William H. 2000. Econometric analysis. 4th ed. New Jersey: Prentice Hall.

[20] Greene, William H. 1981. Sample selection bias as a specification error: A comment. Econometrica 49, no. 3:795-8.

[21] Greenwood, Michael J. 1997. Internal migration in developed countries. In Handbook of population and family economics, ed. Mark R. Rosenzweig and Oded Stark. Amsterdam: Elsevier.

[22] Grant, E. Kenneth, and John Vanderkamp. 1980. The effects of migration on income: A micro study with Canadian data 1965-1971. Canadian Journal of Economics 13, no. 3:381-406.

[23] Ham, John C., Xianghong Li, and Patricia B. Reagan. 2004. Propensity score matching, a distance-based measure of migration, and the wage growth of young men. Unpublished manuscript, The Ohio State University, Ohio.

[24] Heckman, James J. 1976. The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. Annals of Economic and Social Measurement 15:475-92.

[25] Heckman, James J. 1979. Sample selection bias as a specification error. Econometrica 47, no. 1:153-62.

[26] Heckman, James J. 1990. Varieties of selection bias. American Economic Review 80, no. 2:313-18.

[27] Heckman, James J., and Bo E. Honore. 1990. The empirical content of the Roy model. Econometrica 58, no. 8:1121-49.

[28] Heckman, James J, Robert J. Lalonde, and Jeffrey A. Smith. 1999. The economics and econometrics of active labor market programs. In Handbook of labor economics, ed. Orley Ashenfelter and David Card. Amsterdam: Elsevier.

[29] Heiland, Frank. 2004. Trends in East-West German migration from 1989 to 2002. Demographic Research 11, no. 7:173-94.

[30] Hunt, Jennifer. 2000. Why do people still live in East Germany? Working Paper no. 7564, National Bureau of Economic Research, Cambridge, MA.

[31] Hunt, Jennifer. 2001. Post-unification wage growth in East Germany. Review of Economics and Statistics 83:190-5.

[32] Hunt, Jennifer. 2006. Staunching emigration from East Germany: Age and the determinants of migration. Journal of the European Economic Association (forthcoming).

[33] Kornai, János. 1980. Economics of shortage. Amsterdam, New York: North-Holland Publishing Company.

[34] Krueger, Alan B., and Jorn-Steffen Pischke. 1995. A comparison of East and West German labor markets before and after unification. In Differences and changes in wage structures, ed. Richard B. Freeman and Lawrence F. Katz. Chicago, IL: University of Chicago Press.

[35] Nakosteen, Robert A., and Michael A. Zimmer. 1984. Migration and income: The question of self-selection. Southern Economic Journal 46, no. 3:840-51.

[36] Newey, Whitney K. 1984. A method of moments interpretation of sequential estimators. Economics Letters 14:201-6.

[37] Newey, Whitney K. 1985. Maximum likelihood specification testing and conditional moment tests. Econometrica 53, no. 5:1047-70.

[38] Newey, Whitney K. 1988. Two steps series estimation of sample selection models. Unpublished manuscript, Massachusetts Institute of Technology, Cambridge, MA.

[39] OECD. 2001. Economic surveys Germany. Paris: Organisation for Economic Co-operation and Development Press.

[40] Pagan, Adrian. 1984. Econometric issues in the analysis of regressions with generated regressors. International Economic Review 25, no. 1:221-47.

[41] Pagan, Adrian, and Francis G. Vella. 1989. Diagnostic tests for models based on individual data: A survey. Journal of Applied Econometrics 4, suppl.:S29-S59.

[42] Parikh, Ashok, and Michiel van Leuvensteijn. 2003. Internal migration in the regions of Germany: A panel data analysis. Applied Economics Quarterly 49, no. 2:173-92.

[43] Rodgers, Joan R., and John L. Rodgers. 2000. The effect of geographic mobility on male labor force participation in the United States. Journal of Labor Research 21:117-32.

[44] Roy, Andrew D. 1951. Some thoughts on the distribution of earnings. Oxford Economic Papers 3:135-46.

[45] Siebern, Frank. 2000. Better LATE? Instrumental variables estimation of the returns to job mobility during transition. German Economic Review 1, no. 3:335-62.

[46] Sjaastad, Larry. 1961. The costs and returns of human migration. Journal of Political Economy 70:80-93.

[47] SOEP Group. 2001. The German socio-economic panel (SOEP) after more than 15 years - Overview. In Proceedings of the 2000 fourth international conference of German socio-economic panel study users (GSOEP2000), ed. Elke Holst, Dean R. Lillard and Thomas A. DiPrete. Vierteljahrshefte zur Wirtschaftsforschung, Jg 70, Nr. 1, S. 7-14.

[48] Stock, James H., Jonathan H. Wright, and Motohiro Yogo. 2002. A survey of weak instruments and weak identification in generalized method of moments. Journal of Business and Economic Statistics 20, no. 4:518-29.

[49] Tunali, Insan. 2000. Rationality of Migration. International Economic Review 41, no. 4:893-920.

[50] Yankow, Jeffrey J. 2003. Migration, job change, and wage growth: A new perspective on the pecuniary return to geographic mobility. Journal of Regional Science 43, no. 3:483-516.

[51] Yashiv, Eran. 2004. The self-selection of migrant workers revisited. Discussion Paper no. 1094, Instituts zur Zukunft der Arbeit (IZA), Bonn, Germany.

8 Appendix

Table A1: Descriptive statistics

	Migration		Commuting	
	Migrants	Stayers	Commuters	Stayers
mean total annual income	39754	31125	43128	30009
	(26828)	(16937)	(22084)	(16739)
home owner in 1990	0.16	0.33		
border with West Germany in 1990			0.48	0.27
sex	0.42	0.52	0.65	0.49
age in 1990	26.08	31.93	28.59	32.05
	(11.36)	(11.53)	(11.07)	(11.67)
spouse in 1990	0.61	0.74	0.69	0.74
university degree in 1990	0.16	0.09	0.13	0.09
any vocational education in 1990	0.78	0.88	0.83	0.88
employed in government sector in 1990	0.44	0.33	0.31	0.34
blue-collar employee in 1990	0.26	0.35	0.40	0.33
had a telephone in 1990	0.23	0.23	0.28	0.22
annual individual income in 1990	26358	24164	29378	23276
	(21708)	(12320)	(18130)	(11354)
state's unemployment rate in 1992	10.51	10.49	10.70	10.45
	(1.02)	(0.93)	(1.02)	(0.91)

Note: standard deviations in parentheses. Incomes are inflated by regional CPIs to 2001 and are expressed in DM. Sample size varies with the variables, minimum sample sizes are 3043 observations for migration, and 2953 observations for commuting. Mean total annual income is the sum of average annual labour income (sum of wages, second job and self-employment income) and annual social security benefits (such as unemployment benefits, maternity benefits etc).

Table A2: Means of the covariates by instruments

	Migration		Commuting	
	own90=1	own90=0	bwest90=1	bwest90=0
sex	0.53	0.50	0.52	0.51
age in 1990	32.29*	31.21*	31.11	31.82
spouse in 1990	0.76	0.73	0.75	0.73
years of schooling in 1990	11.86*	12.26*	12.27	12.06
university degree in 1990	0.05*	0.11*	0.09	0.09
any vocational education in 1990	0.89	0.87	0.88	0.87
employed in government sector in 1990	0.28*	0.37*	0.36	0.33
blue-collar employee in 1990	0.31*	0.36*	0.33	0.35
had a telephone in 1990	0.24	0.23	0.31*	0.20*
annual individual income in 1990	22758*	24973*	24576	23849

Notes: * difference in means significant at 5%. owner90 is a dummy which equals one if a person owned a house before unification; bwest90 - a dummy which equals one if a person lived in the county that had a common border with West Germany before unification. Incomes are inflated by regional CPIs to 2001 and are expressed in DM.

Table A3: Reduced form estimates

	Migration		Commuting	
	Probit	LPM	Probit	LPM
constant	-0.79	0.18	-3.44	-0.36
	(0.589)	(0.070)	(0.446)	(0.104)
home owner in 1990	-0.42	-0.04		
	(0.093)	(0.008)		
border with West Germany in 1990			0.47	0.11
			(0.061)	(0.015)
sex	-0.15	-0.02	0.35	0.07
	(0.079)	(0.009)	(0.062)	(0.013)
age	-0.02	-0.004	0.07	0.01
	(0.025)	(0.003)	(0.019)	(0.004)
age^2	0.00001	0.00003	-0.001	-0.0002
	(0.0004)	(0.00004)	(0.0003)	(0.0001)
spouse in 1990	-0.15	-0.02	-0.09	-0.02
	(0.103)	(0.012)	(0.081)	(0.018)
university degree in 1990	0.20	0.02	0.25	0.05
	(0.170)	(0.022)	(0.135)	(0.031)
any vocational education in 1990	-0.11	-0.02	-0.15	-0.04
	(0.141)	(0.019)	(0.115)	(0.026)
employed in government sector in 1990	0.12	0.01	-0.05	-0.01
	(0.093)	(0.010)	(0.074)	(0.015)
blue-collar employee in 1990	-0.1	-0.01	0.14	0.03
	(0.100)	(0.009)	(0.071)	(0.015)
unemployment rate in the state, 1992	0.01	-0.001	0.12	0.03
	(0.041)	(0.005)	(0.031)	(0.007)
R^2	0.06	0.03	0.08	0.06
F-test on instrument	30.23		62.52	
# observations	3043		2953	

Note: robust standard errors are in parenthesis. Dependent variable is migrating (columns 1-2) or commuting (columns 3-4) to West Germany. Probit reports coefficients from probit Maximum Likelihood estimation, LPM reports coefficients from linear probability model. Covariates also include dummies for missing 1990 information.

Table A4: Heckman's second stage estimates

	Migration		Commuting	
	Migrants	Stayers	Commuter	Stayers
constant	6.02	6.61	8.70	6.45
	(1.286)	(0.231)	(0.810)	(0.252)
sex	0.74	0.38	0.44	0.38
	(0.125)	(0.022)	(0.064)	(0.027)
age	0.11	0.14	0.06	0.15
	(0.049)	(0.009)	(0.025)	(0.010)
age^2	-0.001	-0.001	-0.0006	-0.001
	(0.0006)	(0.0001)	(0.0003)	(0.0001)
spouse in 1990	-0.35	-0.08	-0.06	-0.08
	(0.157)	(0.028)	(0.068)	(0.029)
state's unemployment rate, 1992	0.07	0.01	0.04	0.01
	(0.055)	(0.010)	(0.030)	(0.013)
university degree in 1990	0.57	0.49	0.47	0.49
	(0.221)	(0.046)	(0.098)	(0.048)
any vocational education in 1990	-0.13	0.13	0.06	0.15
	(0.197)	(0.038)	(0.091)	(0.041)
in government sector in 1990	0.12	0.19	-0.01	0.21
	(0.141)	(0.024)	(0.061)	(0.025)
blue-collar employee in 1990	-0.07	-0.10	0.01	-0.10
	(0.158)	(0.023)	(0.064)	(0.025)
λ	0.67	0.25	-0.02	0.08
	(0.413)	(0.259)	(0.134)	(0.130)
# observations	178	2865	430	2523
CM test 3rd moment	-0.00004		-0.0040	
	(0.0008)		(0.0020)	
CM test 4th moment	0.0005		0.0115	
	(0.0039)		(0.0057)	

Note: standard errors are corrected for heteroskedasticity and for the first step generated regressors, and are reported in parentheses. Depended variable is log of the total annual average income. λ is the inverse Mills ratio. Covariates also include dummies for missing 1990 information. CM test refers to the conditional moment test for normality, in which coefficients (and standard errors) are reported from the regression of 3rd and 4th moments on a constant and scores from probit.

Table A5: Nonparametric second stage estimates

	Migration		Commuting	
	Migrants	Stayers	Commuters	Stayers
constant	7.95	6.63	8.12	6.44
constant_heck	7.97	6.63	8.14	6.45
constant_andr	7.95	6.63	8.14	6.45
	(1.399)	(0.266)	(0.592)	(0.252)
sex	0.72	0.37	0.46	0.37
	(0.125)	(0.024)	(0.060)	(0.022)
age	0.10	0.14	0.06	0.15
	(0.054)	(0.010)	(0.025)	(0.010)
age^2	-0.001	-0.001	-0.001	-0.002
	(0.0006)	(0.0001)	(0.0003)	(0.0001)
spouse in 1990	-0.35	-0.07	-0.07	-0.07
	(0.151)	(0.029)	(0.072)	(0.028)
state's unemployment rate, 1992	0.07	0.01	0.04	0.01
	(0.056)	(0.011)	(0.026)	(0.012)
university degree in 1990	0.53	0.47	0.48	0.46
	(0.240)	(0.045)	(0.082)	(0.044)
any vocational education in 1990	-0.21	0.13	0.07	0.13
	(0.192)	(0.048)	(0.079)	(0.043)
in government sector in 1990	0.12	0.18	0.003	0.22
	(0.142)	(0.023)	(0.061)	(0.023)
blue-collar employee in 1990	-0.06	-0.09	0.002	-0.09
	(0.124)	(0.024)	(0.058)	(0.023)
pscore	-5.86	-0.87	3.88	
	(3.668)	(3.018)	(1.997)	
pscore^2		52.05	-9.54	
		(56.98)	(4.814)	
pscore^3		-378.62		
		(305.96)		
# observations	177	2663	428	2431

Note: standard errors are calculated as in Das et al (2003) and are in paretheses. Depended variable is log of the total annual average income. Constant_heck and constant_andr are intercepts estimated by Heckman (1990) and Andrews and Schafgans (1998) semiparametric procedures respectively. Pscore is the estimated in the first stage propensity to move West. Covariates also include dummy for missing 1990 information.

Table A6: Leave-one-out cross validation

	Migration		Commuting	
pscore order	Migrants	Stayers	Commuters	Stayers
0	84.75	709.03	107.10	643.63
1	84.40	709.57	107.57	643.96
2	85.30	706.88	106.42	644.33
3	86.11	706.49	106.50	644.85
4	85.75	706.51	107.38	644.41
5	85.60	707.12	108.81	644.88

Note: leave-one-out cross-validation criterion is the sum of squares of forecast errors, where all other observations were used to predict each single observation, and the specification with the smallest sum of forecast errors is chosen. Pscore is the estimated in the first stage propensity to move West. Covariates also include dummies for missing 1990 information.

Table A7: Intentions to treat effects, IV (LATE) and OLS estimates

	Intentions to treat: Move	Intentions to treat: Income	Structural IV estimates	OLS estimates
	(1)	(2)	(3)	(4)
		A: Migration		
home owner in 1990	-0.039	0.011		
	(0.008)	(0.020)		
migrate			-0.273	0.305
			(0.538)	(0.055)
		B: Commuting		
border with the West in 1990	0.111	0.022		
	(0.015)	(0.022)		
commute			0.199	0.396
			(0.194)	(0.028)

Note: robust standard errors are in parentheses. Upper panel A shows the estimates for migration, lower panel B - for commuting. Dependent variable in column 1 is migration or commuting dummy respectively, dependent variable in columns 2, 3, 4 is the log of average total annual income. Covariates include gender, age and its square, spouse indicator in 1990, educational and occupational dummies in 1990, state's unemployment rate in 1992 and dummies for missing 1990 information.

Table A8: Treatment effects for movers

Migration			Commuting		
Parametric	Nonparametric	LATE	Parametric	Nonparametric	LATE
-0.19	0.40	-0.27	0.27	0.42	0.20
(0.531)	(0.233)	(0.538)	(0.230)	(0.029)	(0.194)

Note: standard errors are in parentheses. Standard errors of the effects for sample selection models are calculated as for the Oaxaca decomposition (Greene, 2000). Treatment effects are calculated as shown in Section 4. Dependent variable in the regressions is average annual total income. "Parametric" refers to the Heckman's (1976, 1979) two-stage sample selection model; "Nonparametric" - nonparametric sample selection model of Das et al (2003);" LATE" - local average treatment effect of Angrist et al (1996). In the reported nonparametric model the intercept is estimated according to Andrews and Schafgans (1998).

Table A9: Robustness checks

Migration			Commuting		
Parametric	Nonparametric	LATE	Parametric	Nonparametric	LATE
including unemployment in 1990					
-0.13	0.47	-0.25	0.31	0.42	0.23
(0.525)	(0.269)	(0.535)	(0.227)	(0.029)	(0.191)
including household income in 1990					
-0.04	0.32	-0.03	0.27	0.40	0.23
(0.521)	(2.461)	(0.524)	(0.228)	(0.029)	(0.192)
including telephone in 1990					
-0.16	-0.14	-0.24	0.17	0.39	0.10
(0.529)	(1.167)	(0.535)	(0.236)	(0.072)	(0.202)
including political interests in 1990					
-0.39	0.41	-0.38	0.27	0.41	0.21
(0.534)	(0.224)	(0.541)	(0.228)	(0.029)	(0.193)
excluding self-employed					
-0.26	0.07	-0.32	0.31	0.45	0.22
(0.598)	(0.104)	(0.608)	(0.235)	(0.030)	(0.197)
retaining return and multiple migrants					
-0.22	0.07	-0.31	0.30	0.40	0.25
(0.482)	(0.090)	(0.497)	(0.229)	(0.027)	(0.192)
excluding "movers" from the control groups					
0.14	-0.52	0.03	0.24	0.43	0.19
(0.523)	(0.612)	(0.524)	(0.227)	(0.065)	(0.192)

Note: see footnote of Table A8. Results are from the extended model that includes lagged university and vocational education, dummies for government sector and blue-collar employees.

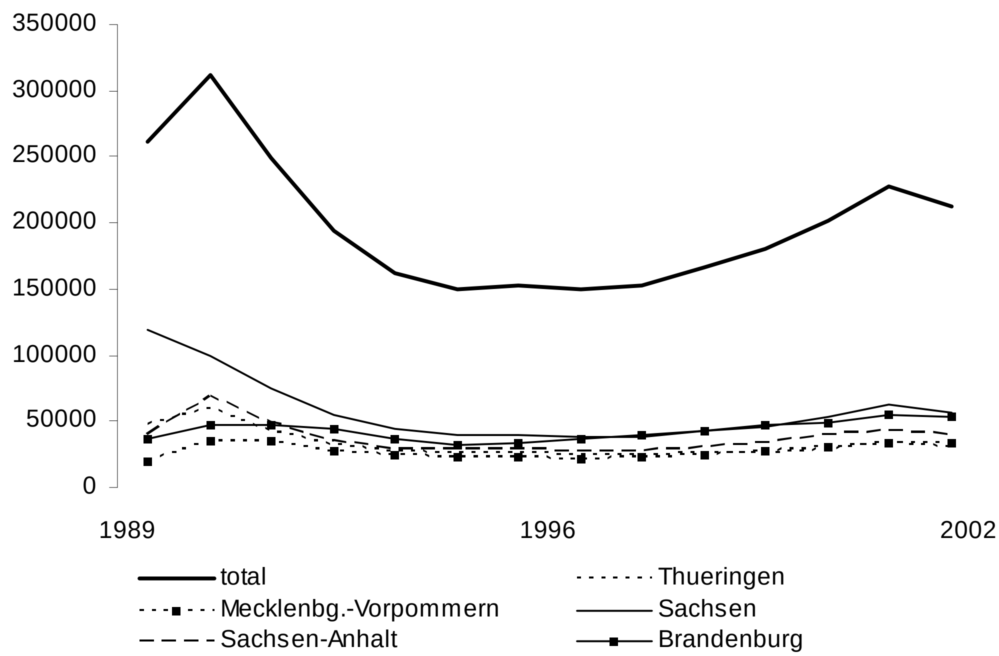

Figure 1: Emigration from East German länder to West Germany after the fall of the Berlin Wall. Source: numbers are from Heiland (2004). Note: East Berlin is omitted due to data unavailability.

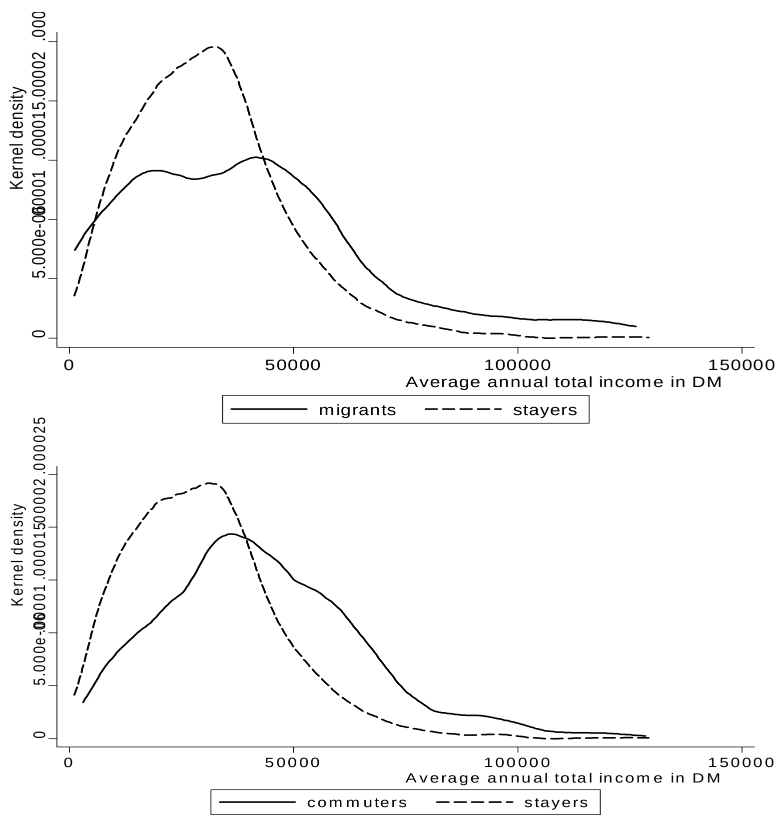

Figure 2: Kernel densities of the average annual total income for movers and stayers in Germany after unification. Source: GSOEP. Notes: see Section 2 for definitions.